AMENDMENTS TO THE CLAIMS

1-2. (CANCELLED)

3. (CURRENTLY AMENDED) A computer-implemented method of constructing a model for predicting whether a test molecule has a chemical or biological property of interest ~~molecular behavior~~ using marker molecules that are known to possess said chemical or biological property of interest, said method comprising:

classifying respective molecules in a training set of reference molecules as either possessing or not possessing ~~at least one~~said chemical or biological property of interest, wherein said classifying is based on experimental data;

selecting, from said training set, a plurality of molecules that possess ~~said at least one~~ the chemical or biological property of interest as target molecules for potential selection as marker molecules for said model;

selecting some of said target molecules as marker molecules for said model by evaluating the predictive accuracy of said potential marker molecules, wherein said evaluating comprises:

computing a numerical value defining a measure of molecular structural similarity for each pair of molecules in said training set using a pre-defined structural similarity metric;

selecting one of said target molecules (T);

sorting all training set molecules in descending order of structural similarity to molecule T as defined by the computed numerical values;

defining, for a first one of said sorted training set molecules (M) a fractions-correctly-predicted metric as a ratio of A/B, wherein B is defined as the total number of training set molecules that have a computed numerical structural similarity to molecule T that is as large or larger than the computed numerical structural similarity between molecules T and M, and wherein A is defined as the number of training set molecules that both (1) have a computed numerical structural similarity with molecule T that is as large or larger than the computed

numerical structural similarity between molecules T and M, and (2) possess the ~~at least one~~ chemical or biological property of interest;

repeating the defining step for other sorted training set molecules;

choosing molecule T as a marker molecule if said number B and said ratio A/B are both above respective threshold values when computed during at least one of said defining steps; and

outputting data indicating that molecule T has been chosen as a marker molecule.

4. (CANCELED)

5. (CANCELED)

6. (PREVIOUSLY PRESENTED) The method of Claim 3, additionally comprising repeating said choosing for a plurality of different threshold values.

7. (PREVIOUSLY PRESENTED) The method of Claim 3, comprising repeating said selecting a target molecule, sorting, defining and choosing steps for other molecules that possess the at least one chemical or biological property at a plurality of different threshold values so as to select a plurality of preliminary sets of marker molecules.

8. (PREVIOUSLY PRESENTED) The method of Claim 7, comprising choosing a final set of marker molecules by making molecular behavior predictions for all molecules in said training set using each one of said preliminary sets of marker molecules, and choosing as said final set of marker molecules the preliminary set that most accurately predicts molecular behavior of molecules of said training set.

9-18. (CANCELED)

19. (PREVIOUSLY PRESENTED) The method of Claim 3, wherein said threshold for B is 5, and said threshold for A/B is 1.